

Klasifikasi *Multi-Label* Postur Tubuh Berbasis MediaPipe dengan Perbandingan Model *Machine Learning*

Fridy Mandita^{1*}, Abdul Azizur Rokhman²

^{1,2}Fakultas Teknologi Elektro dan Informatika Cerdas, Universitas 17 Agustus 1945 Surabaya,
Semolowaru 45 Surabaya

E-mail: ¹fridymandita@untag-sby.ac.id, ²azizrokhman88@gmail.com

*Penulis korespondensi

ABSTRAK

Penelitian ini mengembangkan sistem klasifikasi multi-label untuk mengenali indikasi *Forward Head Posture* (FHP), *Postural Kyphosis* (PK), *Rounded Shoulder Posture* (RSP), dan *Normal Posture* (NP) melalui analisis citra postur statis berdiri yang dapat merepresentasikan perubahan postural berkaitan dengan gaya hidup sedentari. Meskipun demikian, pengukuran dilakukan melalui analisis citra postur statis dalam posisi berdiri untuk mengamati keselarasan kepala, bahu, tulang belakang, dan panggul. MediaPipe Pose digunakan untuk mengekstraksi 33 *landmark* tubuh dari setiap citra. Dataset awal terdiri atas 933 citra yang diperoleh dari dataset publik POLAR dan hasil kurasi mandiri. Setelah pemeriksaan kualitas citra dan kelengkapan *landmark*, sebanyak 54 citra dikeluarkan sehingga diperoleh 879 citra valid. Data dibagi menggunakan *Iterative Stratified Split* dengan rasio 80:20 untuk mempertahankan distribusi kombinasi label. Sebanyak 132 atribut mentah hasil ekstraksi *landmark* dan 54 fitur hasil *feature engineering* membentuk 186 fitur awal. Seleksi fitur menggunakan *Mutual Information* pada data latih menghasilkan 60 fitur terpilih. Model yang dibandingkan meliputi *Light Gradient Boosting Machine*, *Support Vector Machine*, dan *Artificial Neural Network*. Hasil pengujian menunjukkan bahwa SVM memberikan performa terbaik dengan *F1-score macro* sebesar 0,76, *F1-score micro* sebesar 0,79, dan *Hamming Loss* sebesar 0,19. FHP menjadi kelas dengan performa tertinggi, sedangkan PK menjadi kelas yang paling sulit dikenali karena keterbatasan *landmark* dua dimensi pada area vertebra torakal. Hasil penelitian menunjukkan bahwa kombinasi MediaPipe Pose, *feature engineering*, dan *machine learning* berpotensi mendukung pengembangan sistem pemantauan indikasi postur secara waktu nyata.

Article Info

Kata Kunci:

Pose estimation

Postur tubuh

Feature engineering

Komparasi model

Multi-label classification

Riwayat artikel:

Submit 14 Juli 2026

Revisi 18 Juli 2026

Diterima 22 Juli 2026

1. PENDAHULUAN

Perkembangan teknologi dan transformasi digital telah membawa perubahan besar dalam cara manusia bekerja, belajar, dan menjalani aktivitas sehari-hari. Aktivitas yang sebelumnya banyak melibatkan gerak fisik kini semakin sering dilakukan di depan layar komputer, baik di lingkungan kerja, sekolah, maupun perguruan tinggi. Pola aktivitas seperti ini membuat banyak orang menghabiskan waktu dalam posisi duduk dalam durasi yang panjang. Apabila hal ini tidak diimbangi dengan aktivitas fisik yang cukup, kebiasaan tersebut dapat berkembang menjadi gaya hidup sedentari yang berisiko menimbulkan gangguan pada sistem otot dan tulang atau *musculoskeletal system* [1], [2]. Dalam jangka panjang, kondisi ini dapat memengaruhi kerja tubuh, terutama pada tulang belakang, leher, bahu, dan otot-otot penopang postur. Kebiasaan yang berlangsung terus-menerus dapat mendorong munculnya perubahan postur seperti *Forward Head Posture* (FHP) [3], *Postural Kyphosis* [4], serta *Rounded Shoulder Posture* (RSP) [5].

Perubahan postur tubuh umumnya tidak terjadi secara tiba-tiba, melainkan berkembang secara perlahan sehingga sering tidak disadari pada tahap awal. Banyak individu baru menyadari adanya masalah postur setelah

merasakan keluhan seperti nyeri leher, bahu terasa tegang, punggung kaku, cepat lelah, atau berkurangnya kenyamanan saat beraktivitas. Dari sisi biomekanik, postur tubuh yang tidak seimbang dapat mengubah distribusi beban tubuh sehingga tekanan pada tulang belakang dan jaringan lunak di sekitarnya menjadi tidak proporsional [1]. Pada kasus FHP, misalnya, posisi kepala yang cenderung maju ke depan dapat meningkatkan beban pada vertebra servikal, membatasi mobilitas leher, dan berkaitan dengan gangguan fungsi pernapasan serta mobilitas dada [6]. Oleh karena itu, deteksi dini terhadap kelainan postur penting dilakukan agar gangguan postur tidak berkembang menjadi keluhan muskuloskeletal yang lebih serius.

Dampak kelainan postur tidak hanya berkaitan dengan rasa nyeri atau ketidaknyamanan fisik, tetapi juga dapat memengaruhi produktivitas kerja, aktivitas harian, dan kualitas hidup seseorang. Kondisi ini menjadi semakin relevan pada era digital karena aktivitas berbasis layar sering dilakukan dalam durasi lama dengan posisi tubuh yang kurang ergonomis [2]. Selain itu, kelainan postur tidak selalu muncul sebagai satu kondisi yang berdiri sendiri. Seseorang dapat mengalami FHP bersamaan dengan RSP, atau memiliki kombinasi kelainan postur lain yang saling berhubungan secara biomekanik. Situasi ini menunjukkan bahwa pendekatan klasifikasi tunggal atau *single-label classification* belum sepenuhnya mampu menggambarkan kondisi nyata, karena satu individu dapat menunjukkan lebih dari satu indikasi kelainan postur secara bersamaan.

Selama ini, deteksi postur secara konvensional umumnya dilakukan oleh dokter, fisioterapis, atau tenaga ahli melalui observasi langsung maupun alat bantu seperti *posture grid*, goniometer, inclinometer, atau metode pengukuran klinis lainnya. Meskipun metode tersebut dapat memberikan hasil yang baik, penerapannya membutuhkan keahlian khusus, waktu pemeriksaan, biaya, serta kondisi pengamatan yang relatif terkontrol. Penilaian berbasis observasi visual juga berpotensi dipengaruhi subjektivitas pemeriksa [7]. Pada bidang ergonomi, kebutuhan terhadap sistem penilaian postur yang objektif juga semakin kuat karena penilaian otomatis dapat membantu mengurangi variasi penilaian antar-pengamat dan mendukung pencegahan gangguan *muskuloskeletal* [8]. Kondisi ini menimbulkan kebutuhan terhadap sistem deteksi postur yang lebih objektif, otomatis, mudah diakses, dan dapat digunakan untuk mendukung pemantauan postur secara mandiri maupun berkelanjutan.

Kemajuan dalam bidang *Computer Vision* (CV) dan *Machine Learning* (ML) membuka peluang besar untuk mengembangkan sistem analisis postur berbasis citra. Salah satu pendekatan yang banyak digunakan adalah *human pose estimation*, yaitu proses mendeteksi titik-titik penting tubuh manusia dari gambar atau video. Beberapa model *pose estimation* seperti MediaPipe Pose, OpenPose, PoseNet, MoveNet, dan BlazePose telah digunakan dalam berbagai konteks analisis gerakan manusia [7]. Di antara pendekatan tersebut, MediaPipe Pose memiliki keunggulan karena mampu mengekstraksi *landmark* tubuh secara efisien, ringan, dan mendukung pemrosesan citra secara *real-time* [9]. Representasi *landmark* tersebut dapat diolah menjadi data numerik yang lebih ringkas dibandingkan citra RGB penuh, sehingga lebih efisien untuk diproses oleh algoritma *machine learning*.

Berbagai penelitian terdahulu menunjukkan bahwa kombinasi *pose estimation* dan *machine learning* efektif digunakan untuk pengenalan pose dan analisis gerakan manusia. Penelitian [10] menggunakan MediaPipe dan model *machine learning* untuk mengklasifikasikan pose latihan dalam lingkungan *Virtual Reality* (VR) dengan akurasi keseluruhan sebesar 94%. Penelitian lain [11] menunjukkan bahwa fitur hasil *pose estimation* yang dikombinasikan dengan algoritma SVM, Random Forest, dan KNN mampu mengklasifikasikan pose yoga dengan performa tinggi. Sementara itu, penelitian [12] memanfaatkan MediaPipe dan LightGBM untuk mengestimasi sudut rotasi bahu dengan korelasi prediksi yang sangat tinggi dan tingkat kesalahan yang rendah. Penelitian [13] juga menunjukkan bahwa fitur *skeleton*, seperti koordinat sendi, sudut, panjang anggota tubuh, dan simetri tubuh, dapat digunakan untuk pengenalan postur berbasis ML. Hasil penelitian [14] juga menegaskan bahwa algoritma ML klasik masih relevan untuk klasifikasi postur *real-time* karena mampu mencapai akurasi tinggi dengan beban komputasi yang relatif ringan. Temuan-temuan tersebut memperkuat bahwa data *landmark* tubuh dapat menjadi representasi yang efektif dalam analisis postur berbasis citra.

Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada aktivitas dinamis seperti olahraga, yoga, rehabilitasi, gerakan tubuh, atau pose dalam lingkungan virtual [10], [11], [15]. Penelitian yang berfokus pada postur statis masih banyak membatasi klasifikasi pada kategori umum seperti postur baik dan buruk, tanpa mengidentifikasi jenis kelainan postur secara spesifik [2]. Selain itu, sebagian besar pendekatan masih menggunakan *single-label classification*, sehingga belum optimal untuk mendeteksi kemungkinan munculnya lebih dari satu kelainan postur pada individu yang sama. Padahal, dalam kondisi nyata, kelainan seperti FHP, PK, dan RSP dapat muncul secara bersamaan karena saling berkaitan dalam rantai biomekanik kepala, bahu, tulang belakang, dan panggul.

Selain kesenjangan pada aspek klasifikasi, terdapat pula kesenjangan pada representasi data. Pendekatan berbasis citra RGB penuh seperti *Convolutional Neural Network* (CNN) mampu mempelajari pola visual secara langsung dan telah menunjukkan performa yang tinggi dalam berbagai permasalahan klasifikasi citra. Penelitian [16], misalnya, menunjukkan bahwa CNN mampu menghasilkan akurasi klasifikasi citra yang tinggi. Namun, penelitian tersebut juga memperlihatkan bahwa peningkatan ukuran citra dan jumlah *epoch* dapat memperpanjang

waktu pelatihan, sedangkan peningkatan akurasi yang diperoleh tidak selalu signifikan. Temuan ini menunjukkan adanya kompromi antara performa klasifikasi dan kebutuhan komputasi pada pendekatan berbasis citra penuh. Selain itu, pendekatan tersebut berpotensi lebih rentan terhadap variasi lingkungan, seperti pencahayaan, latar belakang, tekstur pakaian, dan sudut pengambilan gambar [14], [17].

Sebaliknya, pendekatan berbasis *skeleton* menggunakan *landmark* MediaPipe dapat menyederhanakan citra menjadi representasi koordinat tubuh yang lebih efisien, interpretatif, dan relatif lebih menjaga privasi karena sistem tidak perlu memproses identitas visual secara penuh [7]. Representasi ini juga memungkinkan pengembangan fitur biomekanik spasial, seperti jarak antartitik tubuh, rasio anatomi, pergeseran posisi kepala, bahu, dan panggul, serta sudut atau proksi sudut yang relevan untuk mendeteksi kelainan postur [18].

Selain pemilihan representasi data, keberhasilan sistem klasifikasi juga dipengaruhi oleh relevansi fitur dan kesesuaian algoritma dengan karakteristik data. Penelitian komparatif pada bidang kesehatan menunjukkan bahwa seleksi fitur statistik dapat digunakan untuk mengurangi fitur yang kurang relevan sebelum membangun dan membandingkan beberapa model, seperti SVM, XGBoost, *Random Forest*, *Naive Bayes*, dan ANN [19]. Penelitian tersebut memperoleh XGBoost sebagai model dengan performa terbaik pada data yang digunakan. Namun, penelitian komparatif lainnya yang mengevaluasi KNN, *Naive Bayes*, Regresi Logistik, SVM, dan *Neural Network* menunjukkan bahwa peringkat kinerja model dapat berubah bergantung pada karakteristik data, konfigurasi model, dan metrik evaluasi yang digunakan [20]. Meskipun kedua penelitian tersebut menggunakan data prediksi diabetes, temuan metodologisnya relevan untuk penelitian ini karena menegaskan pentingnya seleksi fitur dan evaluasi beberapa algoritma secara komparatif, bukan menetapkan satu algoritma sebagai model yang secara universal paling unggul.

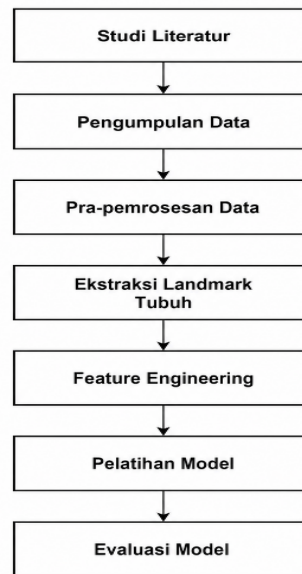
Berdasarkan kesenjangan tersebut, penelitian ini mengembangkan sistem klasifikasi *multi-label* untuk mengenali indikasi permasalahan postur tubuh berbasis *computer vision* dan *machine learning* menggunakan MediaPipe Pose. MediaPipe Pose digunakan untuk mengekstraksi *landmark* tubuh dari citra postur statis manusia dalam posisi berdiri. *Landmark* tersebut kemudian diolah menjadi fitur spasial dan biomekanik yang merepresentasikan hubungan posisi kepala, bahu, tulang belakang, dan panggul. Fitur hasil rekayasa digunakan sebagai masukan untuk membandingkan performa LightGBM, SVM, dan ANN.

Penelitian ini berfokus pada klasifikasi *multi-label* terhadap *Forward Head Posture* (FHP), *Postural Kyphosis* (PK), *Rounded Shoulder Posture* (RSP), dan *Normal Posture* (NP). Pendekatan *multi-label* digunakan karena satu citra dapat menunjukkan lebih dari satu indikasi perubahan postur secara bersamaan. Kebaruan penelitian ini terletak pada penerapan klasifikasi *multi-label* untuk mengenali indikasi postur statis menggunakan fitur spasial dan biomekanik hasil ekstraksi MediaPipe Pose. Selain itu, penelitian ini melakukan evaluasi komparatif terhadap model berbasis *gradient boosting*, model berbasis margin dan kernel, serta model jaringan saraf. Perbandingan tersebut dilakukan untuk memperoleh model yang paling sesuai dengan karakteristik fitur postur yang digunakan, mengingat tidak terdapat satu algoritma yang selalu memberikan performa terbaik pada seluruh jenis data dan permasalahan klasifikasi [19], [20].

Penelitian ini dilatarbelakangi oleh perubahan postur tubuh yang dapat berkaitan dengan gaya hidup sedentari dan kebiasaan duduk berkepanjangan. Meskipun demikian, pengambilan dan analisis data dilakukan menggunakan citra tubuh dalam posisi berdiri. Postur statis berdiri digunakan untuk mengamati keselarasan kepala, bahu, tulang belakang, dan panggul serta mengenali indikasi perubahan postural yang tetap terlihat setelah seseorang melakukan aktivitas sehari-hari. Oleh karena itu, penelitian ini tidak mengevaluasi ergonomi tubuh ketika sedang duduk maupun gerakan tubuh secara dinamis, tetapi berfokus pada klasifikasi *multi-label* terhadap indikasi FHP, PK, RSP, dan NP berdasarkan citra postur statis dalam posisi berdiri. Hasil penelitian diharapkan dapat menjadi dasar pengembangan sistem pemantauan indikasi postur secara waktu nyata serta memperluas penerapan *machine learning* dalam bidang kesehatan digital, ergonomi, rehabilitasi, dan analisis postur tubuh berbasis *computer vision*.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen komputasional untuk mengembangkan sistem klasifikasi kelainan postur tubuh berbasis citra. Sistem dirancang untuk mengklasifikasikan kondisi postur tubuh manusia dalam posisi berdiri menggunakan pendekatan klasifikasi *multi-label*. Secara umum, tahapan penelitian meliputi pengumpulan *dataset*, pelabelan data, pra-pemrosesan, ekstraksi *landmark* tubuh menggunakan MediaPipe Pose, rekayasa fitur, seleksi fitur, pelatihan model, dan evaluasi performa model. Alur tahapan penelitian ditunjukkan pada Gambar 1. Tahapan ini disusun berdasarkan proses mulai dari pengumpulan data, pra-pemrosesan, ekstraksi *landmark*, *feature engineering*, pelatihan model, hingga evaluasi.



Gambar 1 Tahapan Penelitian

2.1 Pengambilan Dataset

Dataset yang digunakan berupa citra tubuh manusia dalam posisi berdiri yang merepresentasikan empat kategori postur, yaitu FHP, PK, RSP, dan NP. Data utama diperoleh dari *dataset* publik *Posture-Level Action Recognition* (POLAR), serta dilengkapi dengan citra tambahan dari sumber daring publik yang dikurasi berdasarkan kesesuaian sudut pengambilan, keterlihatan tubuh, kualitas citra, dan relevansi terhadap kategori postur. Citra yang digunakan dibatasi pada data yang tersedia secara publik dan diproses hanya dalam bentuk *landmark* tubuh untuk keperluan analisis postur. Karena satu citra dapat menunjukkan lebih dari satu indikasi kelainan postur, struktur label disusun dalam format *multi-label*, yaitu setiap kelas postur direpresentasikan sebagai label biner.

Tabel 1. Deskripsi *Dataset*

Kategori Postur	Jumlah Data
FHP	676
PK	296
RSP	402
NP	242

Jumlah pada Tabel 1 menunjukkan distribusi label dalam pendekatan *multi-label*, sehingga total label tidak selalu sama dengan jumlah citra unik karena satu citra dapat memiliki lebih dari satu kondisi postur. Contoh citra *dataset* ditampilkan pada Gambar 2 untuk memberikan gambaran variasi postur yang dianalisis dalam penelitian ini.



Gambar 2. (a) *Forward Head Posture* (FHP) (b) FHP dan *Postural Kyphosis* (PK)

2.2 Labelling Data

Pelabelan data dilakukan secara manual dengan mengacu pada indikator postural berdasarkan literatur dan inspeksi visual terhadap posisi tubuh. Proses pelabelan dilakukan melalui dua tahap, yaitu inspeksi visual dan validasi kuantitatif berbasis proksi *landmark* MediaPipe. Inspeksi visual dilakukan dengan menilai posisi kepala terhadap bahu dan panggul untuk mendeteksi FHP, kelengkungan area toraks untuk mendeteksi PK, serta posisi bahu yang maju terhadap panggul untuk mendeteksi RSP. Validasi kuantitatif digunakan untuk menjaga konsistensi label berdasarkan indikator seperti offset horizontal telinga terhadap panggul, *Curvature Index* (CI), dan offset horizontal bahu terhadap panggul. Ringkasan kriteria pelabelan ditunjukkan pada Tabel 2.

Tabel 2. Kriteria Kuantitatif Pelabelan

Kondisi	Acuan Klinis	Proksi MediaPipe	Kriteria Label Positif
FHP	$CVA < 50^\circ$	ear_hip_x_norm	Kepala tampak maju di depan garis vertikal panggul dan dikonfirmasi oleh offset telinga-panggul yang tinggi
PK	$Cobb\ Angle > 40^\circ$	Curvature Index	Punggung atas tampak membulat dan dikonfirmasi oleh $Curvature\ Index > 1,0$
RSP	$FSA > 52^\circ$	rsp_shoulder_hip_x_norm	Bahu tampak maju terhadap garis vertikal panggul
NP	Tidak ada deviasi signifikan	Semua proksi dalam rentang normal	Tidak memenuhi kriteria positif untuk FHP, PK, maupun RSP

Untuk menghindari kesalahan pelabelan, dilakukan validasi kontradiksi label. Pada pendekatan *multi-label*, citra dapat memiliki kombinasi label seperti FHP dan RSP, tetapi tidak boleh memiliki label NP bersamaan dengan label kelainan postur. Oleh karena itu, *dataset* diperiksa menggunakan aturan logika bahwa apabila label NP bernilai 1, maka label FHP, PK dan RSP harus bernilai 0

2.3 Pra-Pemrosesan Data dan Augmentasi

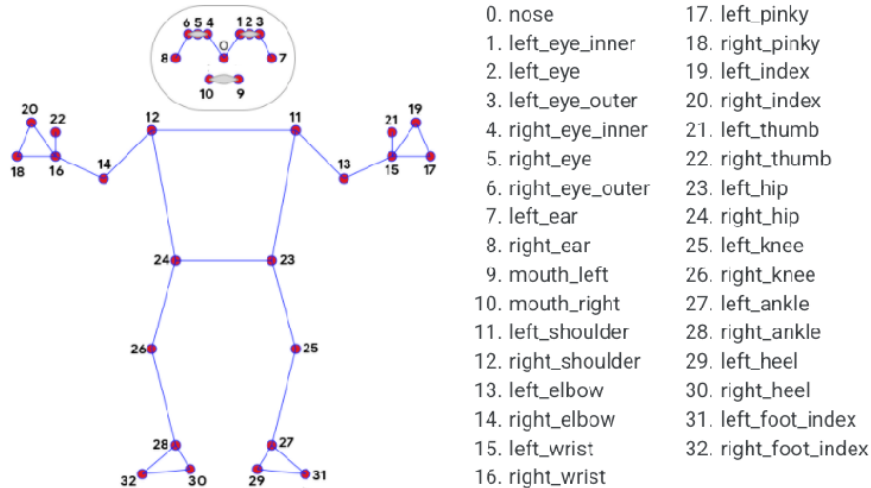
Tahap pra-pemrosesan dilakukan untuk memastikan bahwa data citra dan fitur yang dihasilkan memiliki format yang konsisten sebelum digunakan dalam pelatihan model. Citra yang tidak dapat menghasilkan *landmark* tubuh secara lengkap atau memiliki kualitas deteksi yang tidak memenuhi kriteria dikeluarkan dari dataset. Berdasarkan proses validasi tersebut, diperoleh 879 citra valid dari total awal sebanyak 933 citra.

Augmentasi data diterapkan untuk meningkatkan variasi data latih dan mengurangi risiko *overfitting*. Teknik augmentasi yang digunakan meliputi penambahan *Gaussian noise* dengan nilai simpangan baku $\sigma = 0,01$, perubahan skala tubuh berupa pembesaran dan pengecilan, serta translasi acak pada koordinat *landmark* dengan rentang $\pm 0,02$. Proses tersebut menghasilkan tiga variasi tambahan untuk setiap data latih.

Augmentasi hanya diterapkan pada data latih setelah proses pembagian dataset dilakukan. Data uji tidak mengalami proses augmentasi dan tetap menggunakan data asli. Ketentuan ini diterapkan agar evaluasi model menggambarkan kemampuan generalisasi terhadap data yang belum pernah digunakan maupun dimodifikasi selama proses pelatihan.

2.4 Proses Ekstraksi Landmark

Ekstraksi titik tubuh dilakukan menggunakan MediaPipe Pose. Sebelum diproses, citra dikonversi dari format BGR ke RGB dan diberikan peningkatan kontras ringan untuk membantu deteksi pada kondisi pencahayaan yang bervariasi. MediaPipe Pose mengekstraksi 33 titik *landmark* tubuh yang merepresentasikan bagian tubuh seperti kepala, bahu, siku, panggul, dan lutut. Setiap *landmark* terdiri atas empat atribut, yaitu koordinat x, y, z, dan *visibility*, sehingga setiap citra menghasilkan 132 nilai koordinat mentah. Ilustrasi hasil ekstraksi *landmark* MediaPipe Pose diilustrasikan pada Gambar 3.



Gambar 3. MediaPipe Pose Landmark

Koordinat mentah hasil ekstraksi MediaPipe tidak langsung digunakan sebagai masukan model karena masih dipengaruhi oleh skala citra, ukuran tubuh subjek, serta jarak subjek terhadap kamera. Oleh karena itu, dilakukan rekayasa fitur atau *feature engineering* untuk mengubah koordinat mentah menjadi fitur biomekanik yang lebih bermakna. Tahap awal dilakukan dengan menghitung titik tengah organ bilateral, seperti telinga, bahu, dan panggul. Perhitungan titik tengah dilakukan menggunakan Persamaan (1).

$$center = \frac{left_{point} + right_{point}}{2} \quad (1)$$

Dengan:

$center$: titik tengah atau koordinat rata-rata dari dua titik landmark yang berpasangan

$left_{point}$: koordinat titik landmark pada sisi kiri tubuh, misalnya telinga kiri, bahu kiri, atau panggul kiri.

$right_{point}$: koordinat titik landmark pada sisi kanan tubuh, misalnya telinga kanan, bahu kanan, atau panggul kanan

Persamaan tersebut digunakan untuk menghitung titik tengah dari dua titik landmark bilateral, yaitu titik tubuh bagian kiri dan kanan. Nilai $center$ diperoleh dengan menjumlahkan koordinat $left_{point}$ dan $right_{point}$, kemudian membaginya dengan dua. Hasil perhitungan ini digunakan sebagai representasi titik tengah organ tubuh, seperti telinga, bahu, atau panggul, agar posisi tubuh dapat dianalisis secara lebih stabil.

Selanjutnya, panjang torso digunakan sebagai basis normalisasi agar fitur jarak tidak bergantung pada ukuran tubuh atau jarak subjek terhadap kamera. Panjang torso dihitung menggunakan jarak Euclidean antara titik tengah bahu dan titik tengah panggul sebagaimana ditunjukkan pada Persamaan (2).

$$torso_length = \sqrt{(x_{shoulder} - x_{hip})^2 + (y_{shoulder} - y_{hip})^2} \quad (2)$$

Dengan:

$torso_length$: panjang torso atau jarak antara titik tengah bahu dan titik tengah panggul

$x_{shoulder}$: koordinat x dari titik tengah bahu

$y_{shoulder}$: koordinat y dari titik tengah bahu

x_{hip} : koordinat x dari titik tengah panggul

y_{hip} : koordinat y dari titik tengah panggul

Persamaan tersebut digunakan untuk menghitung panjang torso sebagai jarak Euclidean antara titik tengah bahu dan titik tengah panggul. Nilai $torso_length$ diperoleh dari selisih koordinat horizontal dan vertikal antara bahu dan panggul, kemudian dihitung menggunakan akar dari jumlah kuadrat kedua selisih tersebut. Dalam penelitian ini,

torso_length digunakan sebagai dasar normalisasi fitur jarak agar nilai fitur tidak bergantung pada ukuran tubuh subjek maupun jarak subjek terhadap kamera.

Fitur untuk FHP diperoleh dengan mengukur pergeseran horizontal telinga terhadap panggul yang telah dinormalisasi menggunakan panjang torso. Fitur ini digunakan sebagai proksi terhadap *Craniovertebral Angle* (CVA). Contoh perhitungan numerik fitur *ear_hip_x_norm* ditunjukkan pada Tabel 3.

Tabel 3. Contoh Perhitungan Fitur

Tahap	Parameter	Nilai
Telinga kiri	Koordinat x	0,35
Telinga kanan	Koordinat x	0,36
Panggul kiri	Koordinat x	0,50
Panggul kanan	Koordinat x	0,51
<i>ear_center x</i>	$(0,35 + 0,36) / 2$	0,355
<i>hip_center x</i>	$(0,50 + 0,51) / 2$	0,505
Selisih absolut	$ 0,355 - 0,505 $	0,150
Panjang torso	Basis normalisasi	0,350
<i>ear_hip x norm</i>	$0,150 / 0,350$	0,428

Selain FHP, penelitian ini juga menghitung fitur untuk PK menggunakan CI. Indeks ini membandingkan panjang jalur segmen telinga–bahu dan bahu–panggul terhadap jarak langsung telinga–panggul. Jika nilai indeks semakin besar dari 1,0, maka kelengkungan tubuh semakin tinggi. Perhitungan CI ditunjukkan pada Persamaan (3).

$$CI = \frac{d_{ear,shoulder} + d_{shoulder,hip}}{d_{ear,hip}} \quad (3)$$

Dengan

$d_{ear,shoulder}$: jarak Euclidean antara titik tengah telinga dan bahu,

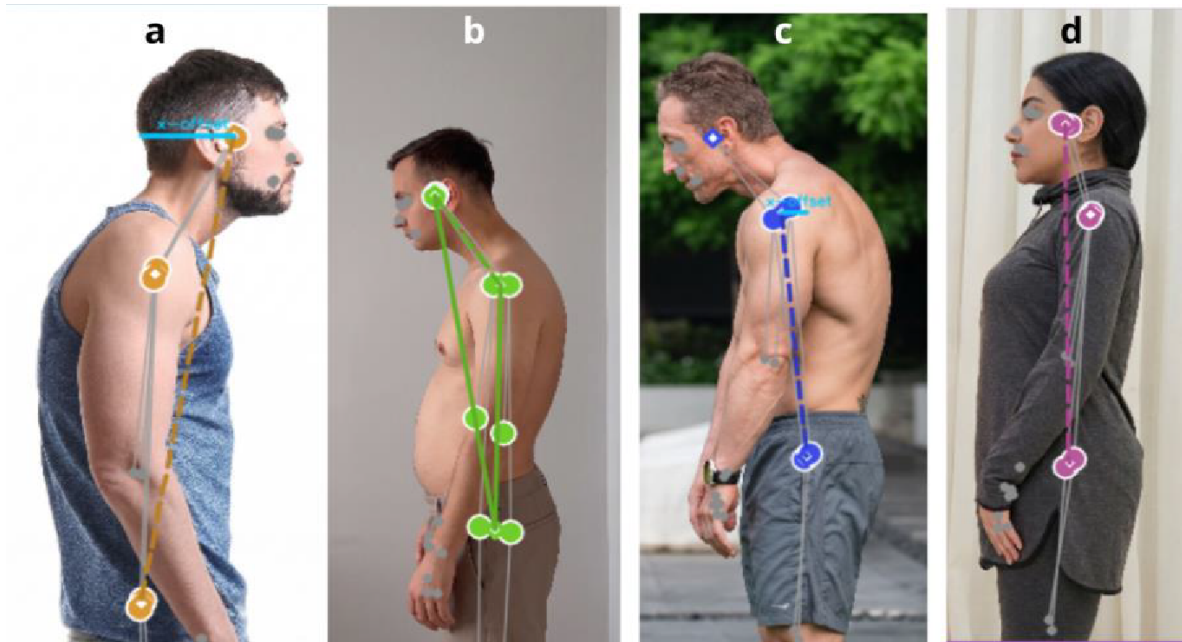
$d_{shoulder,hip}$: jarak Euclidean antara titik tengah bahu dan panggul

$d_{ear,hip}$: jarak langsung antara titik tengah telinga dan panggul.

Persamaan CI digunakan untuk mengukur tingkat kelengkungan tubuh bagian atas berdasarkan hubungan antara titik telinga, bahu, dan panggul. Nilai CI dihitung dengan membandingkan jumlah jarak antara telinga ke bahu dan bahu ke panggul terhadap jarak langsung antara telinga dan panggul. Semakin besar nilai CI dari 1, semakin besar indikasi kelengkungan postur tubuh bagian atas.

2.5 Proses Feature Engineering

Setiap kondisi postur direpresentasikan menggunakan kombinasi titik tubuh yang berbeda. FHP direpresentasikan melalui titik telinga, bahu, dan panggul. PK melibatkan titik telinga, bahu, siku, dan panggul untuk menggambarkan kelengkungan tubuh bagian atas. RSP direpresentasikan terutama melalui titik bahu dan panggul, sedangkan NP dianalisis berdasarkan keselarasan relatif antara titik telinga, bahu, dan panggul. Visualisasi titik kunci yang digunakan pada setiap kelas postur ditampilkan pada Gambar 4.



Gambar 4. Keypoints a) FHP, b) PK, c) RSP, d) NP

Pada Gambar 4, mempresentasikan kombinasi titik tubuh yang digunakan untuk merepresentasikan setiap kondisi postur. Pada Gambar 4(a), FHP dianalisis berdasarkan posisi telinga, bahu, dan panggul. Pergeseran horizontal antara telinga dan panggul digunakan sebagai dasar perhitungan fitur *ear_hip_x_norm*, yang berperan sebagai proksi terhadap *Craniovertebral Angle* (CVA). Gambar 4(b) memperlihatkan representasi PK dengan melibatkan titik telinga, bahu, siku, dan panggul untuk menangkap pola kelengkungan tubuh bagian atas melalui *Curvature Index* (CI). Selanjutnya, Gambar 4(c) menunjukkan representasi RSP berdasarkan pergeseran horizontal bahu terhadap panggul melalui fitur *shoulder_hip_x_norm*, yang digunakan sebagai proksi terhadap *Forward Shoulder Angle* (FSA). Sementara itu, Gambar 4(d) memperlihatkan NP sebagai kondisi ketika titik telinga, bahu, dan panggul relatif sejajar secara vertikal tanpa deviasi horizontal yang signifikan.

Secara keseluruhan, proses rekayasa fitur menghasilkan 54 fitur turunan yang dikelompokkan ke dalam enam blok analisis. Blok tersebut terdiri atas indikator FHP, indikator PK, indikator RSP, indikator keselarasan global tubuh, indikator kualitas dan nilai *visibility landmark*, serta indikator NP. Ringkasan jumlah fitur dan fokus analisis pada setiap blok disajikan pada Tabel 4.

Tabel 4. Ringkasan Blok Fitur

Blok	Indikator	Jumlah Fitur	Fokus Analisis
A	FHP	12	Pergeseran kepala, sudut leher, dan proksi CVA
B	PK	16	Kelengkungan tubuh, <i>Curvature Index</i> , dan rasio segmen
C	RSP	9	Pergeseran bahu terhadap panggul dan proksi FSA
D	Keselarasan global	7	Deviasi global, kemiringan tubuh, dan pola kombinasi <i>multi-label</i>
E	Kualitas landmark	5	Rata-rata, nilai minimum, dan konsistensi <i>visibility</i>
F	NP	5	Kedekatan terhadap kondisi postur seimbang
Total		54	

2.6 Seleksi Fitur dan Pembagian Data

Dataset valid sebanyak 879 citra dibagi menjadi data latih dan data uji menggunakan metode *Iterative Stratified Split* dengan rasio 80:20. Pendekatan tersebut digunakan untuk mempertahankan distribusi kombinasi label secara proporsional dalam skenario klasifikasi *multi-label*. Hasil pembagian dataset terdiri atas 702 citra sebagai data latih dan 177 citra sebagai data uji.

Ekstraksi menggunakan MediaPipe Pose menghasilkan 33 *landmark* tubuh. Setiap *landmark* terdiri atas koordinat x, y, dan z, serta nilai *visibility*, sehingga setiap citra menghasilkan 132 atribut mentah. Selanjutnya, proses *feature engineering* menghasilkan 54 fitur turunan yang mencakup fitur jarak, sudut, rasio, pergeseran posisi, keselarasan tubuh, dan kualitas *landmark*. Penggabungan 132 atribut mentah dan 54 fitur hasil rekayasa menghasilkan sebanyak 186 fitur awal.

Sebelum proses seleksi fitur dilakukan, terlebih dahulu dilakukan penyaringan terhadap atribut yang dinilai kurang reliabel untuk analisis citra dua dimensi, terutama beberapa atribut koordinat z dan nilai *visibility* mentah. Penyaringan tersebut bertujuan untuk mengurangi penggunaan atribut yang kurang stabil atau kurang representatif dalam menggambarkan kondisi postur berdasarkan citra dua dimensi.

Seleksi fitur kemudian dilakukan menggunakan metode *Mutual Information* (MI) untuk mengukur tingkat keterkaitan antara setiap fitur dan masing-masing label target. Fitur dengan nilai MI yang lebih tinggi dianggap memiliki informasi yang lebih relevan untuk membedakan kondisi postur sehingga digunakan sebagai dasar dalam menentukan fitur yang dipertahankan.

Untuk mencegah terjadinya *data leakage*, seluruh proses penyaringan dan seleksi fitur dilakukan hanya menggunakan data latih. Data uji tidak digunakan dalam perhitungan nilai MI maupun dalam penentuan fitur yang dipertahankan. Daftar fitur yang diperoleh dari proses seleksi pada data latih selanjutnya diterapkan tanpa perubahan pada data uji.

Berdasarkan proses penyaringan dan seleksi tersebut, diperoleh 60 fitur akhir yang digunakan sebagai masukan untuk model LightGBM, SVM, dan ANN. Dengan demikian, 60 fitur akhir tidak hanya berasal dari 54 fitur hasil rekayasa, tetapi dipilih dari keseluruhan 186 fitur awal yang mencakup atribut mentah MediaPipe Pose dan fitur hasil rekayasa.

Standarisasi fitur diterapkan pada model SVM dan ANN karena kedua model tersebut sensitif terhadap perbedaan skala antar fitur. Nilai rata-rata dan standar deviasi untuk setiap fitur dihitung hanya menggunakan data latih. Parameter standarisasi yang diperoleh dari data latih kemudian digunakan untuk mentransformasikan data latih dan data uji. Data uji tidak digunakan dalam perhitungan parameter standarisasi sehingga konsistensi pemrosesan tetap terjaga dan kebocoran informasi dari data uji dapat dicegah.

Tabel 5. Ringkasan Model Klasifikasi

Model	Karakteristik Utama	Strategi <i>Multi-label</i>
LightGBM	Model <i>gradient boosting</i> untuk data tabular berfitur numerik	<i>Binary relevance</i> melalui <i>MultiOutputClassifier</i>
SVM	Model berbasis kernel RBF untuk pola non-linear	<i>Binary relevance</i> melalui <i>MultiOutputClassifier</i>
ANN	Model jaringan saraf dengan keluaran sigmoid untuk setiap label	Keluaran <i>multi-label</i> secara langsung

2.7 Pelatihan Model

Tahap pelatihan model bertujuan untuk memetakan hubungan nonlinier antara 60 fitur numerik terpilih dan empat label kondisi postur, yaitu *Forward Head Posture* (FHP), *Postural Kyphosis* (PK), *Rounded Shoulder Posture* (RSP), dan *Normal Posture* (NP). Karena penelitian ini menggunakan klasifikasi *multi-label*, satu citra dapat memperoleh lebih dari satu label positif secara bersamaan. Oleh karena itu, setiap label diperlakukan sebagai keluaran biner yang menunjukkan ada atau tidaknya indikasi kondisi postur tertentu.

Strategi penanganan klasifikasi *multi-label* disesuaikan dengan karakteristik masing-masing model. LightGBM dan SVM pada dasarnya dirancang untuk menangani satu target keluaran. Oleh karena itu, permasalahan *multi-label* ditransformasikan menjadi empat masalah klasifikasi biner yang independen menggunakan strategi *Binary Relevance* melalui *MultiOutputClassifier*. Dengan pendekatan tersebut, satu model biner dilatih untuk setiap label, yaitu FHP, PK, RSP, dan NP. Sebaliknya, ANN menangani klasifikasi *multi-label* secara langsung melalui empat neuron keluaran dengan fungsi aktivasi *sigmoid*. Setiap neuron menghasilkan probabilitas secara independen untuk satu label sehingga satu citra dapat diklasifikasikan ke dalam lebih dari satu kondisi postur.

LightGBM dipilih karena efisien dalam menangani data tabular yang terdiri atas fitur-fitur numerik. Pada penelitian ini, LightGBM dikonfigurasi menggunakan objektif biner karena setiap label diperlakukan sebagai permasalahan klasifikasi biner independen melalui strategi *Binary Relevance*. Empat estimator LightGBM dilatih secara terpisah untuk memprediksi label FHP, PK, RSP, dan NP. Parameter utama yang digunakan meliputi *n_estimators* sebesar 500, *learning_rate* sebesar 0,05, *max_depth* sebesar 6, *subsample* sebesar 0,8, *colsample_bytree* sebesar 0,8, dan *min_child_samples* sebesar 10.

SVM digunakan sebagai model berbasis *kernel* untuk mempelajari pola nonlinier pada fitur postur. *Kernel Radial Basis Function* (RBF) dipilih karena hubungan antara fitur jarak, sudut, rasio, pergeseran posisi, dan keselarasan tubuh tidak selalu dapat dipisahkan secara linier. Seperti LightGBM, SVM diterapkan menggunakan strategi *Binary Relevance* melalui *MultiOutputClassifier* sehingga empat model SVM biner dilatih secara independen untuk setiap label. Parameter *gamma* diatur menggunakan nilai *scale*, sedangkan parameter *C* ditetapkan sebesar 10. Parameter *class_weight* diatur menjadi *balanced* untuk membantu menangani ketidakseimbangan kelas. Selain itu, parameter *probability* diaktifkan agar model dapat menghasilkan probabilitas prediksi untuk setiap label.

ANN yang digunakan memiliki arsitektur *Multilayer Perceptron* dengan 60 fitur masukan dan tiga lapisan tersembunyi yang masing-masing terdiri atas 128, 64, dan 32 neuron. Fungsi aktivasi ReLU digunakan pada setiap lapisan tersembunyi, sedangkan *Batch Normalization* dan *Dropout* sebesar 0,3 diterapkan untuk menjaga stabilitas proses pelatihan dan mengurangi risiko *overfitting*. Lapisan keluaran terdiri atas empat neuron dengan fungsi aktivasi *sigmoid*. Setiap neuron merepresentasikan probabilitas label FHP, PK, RSP, atau NP secara independen. Oleh karena itu, berbeda dengan LightGBM dan SVM, ANN tidak memerlukan *MultiOutputClassifier* karena keempat label diprediksi secara langsung dalam satu arsitektur jaringan. Model dioptimasi menggunakan Adam dengan *learning rate* sebesar 0,001 dan fungsi kerugian *Binary Cross-Entropy*.

Pada proses pelatihan ANN, sebagian data latih dipisahkan dan digunakan sebagai data validasi internal. Data validasi digunakan untuk memantau perubahan nilai akurasi dan *loss* selama proses pelatihan serta mengidentifikasi kemungkinan terjadinya *overfitting*. Data validasi tersebut berasal dari data latih dan bukan dari 177 citra yang telah ditetapkan sebagai data uji. Dengan demikian, data uji tetap disimpan secara terpisah dan hanya digunakan untuk mengevaluasi performa akhir model setelah seluruh proses pelatihan selesai dilakukan.

Distribusi label pada dataset bersifat tidak seimbang, terutama karena jumlah data berlabel PK lebih sedikit dibandingkan data berlabel FHP. Untuk mengurangi kecenderungan model dalam memprioritaskan kelas mayoritas, diterapkan strategi pembobotan kelas atau *sample weighting*. Pada SVM, penanganan ketidakseimbangan kelas dilakukan melalui parameter *class_weight* yang diatur menjadi *balanced*. Sementara itu, pada LightGBM dan ANN, bobot sampel dimasukkan ke dalam proses pelatihan sehingga kesalahan prediksi pada kelas dengan jumlah data lebih sedikit memperoleh penalti yang lebih besar.

Keluaran model berupa probabilitas prediksi untuk setiap label kemudian dikonversi menjadi label biner menggunakan nilai *threshold* sebesar 0,5. Probabilitas yang lebih besar atau sama dengan 0,5 ditetapkan sebagai label 1, yang menunjukkan bahwa kondisi postur terdeteksi. Sebaliknya, probabilitas kurang dari 0,5 ditetapkan sebagai label 0, yang menunjukkan bahwa kondisi postur tidak terdeteksi. Proses penentuan label dilakukan secara independen untuk FHP, PK, RSP, dan NP sehingga satu citra dapat menghasilkan lebih dari satu label positif secara bersamaan.

2.8 Evaluasi Model

Pengujian dilakukan untuk mengukur kemampuan model *machine learning* dalam melakukan generalisasi terhadap data baru dan mengenali pola kelainan postur tubuh secara akurat. Proses evaluasi menggunakan 177 citra data uji yang tidak dilibatkan dalam proses augmentasi, seleksi fitur, standarisasi parameter, pelatihan, maupun validasi model. Skenario pengujian dirancang untuk membandingkan performa tiga algoritma, yaitu LightGBM, SVM, dan ANN, menggunakan metrik evaluasi yang sesuai untuk klasifikasi *multi-label*.

Setiap model menghasilkan probabilitas prediksi secara independen untuk empat label kondisi postur, yaitu FHP, PK, RSP, dan NP. Sebelum metrik evaluasi dihitung, probabilitas setiap label dikonversi menjadi keluaran biner menggunakan nilai *threshold* sebesar 0,5. Probabilitas yang lebih besar atau sama dengan 0,5 ditetapkan sebagai label 1, yang menunjukkan bahwa kondisi postur terdeteksi. Sebaliknya, probabilitas kurang dari 0,5 ditetapkan sebagai label 0, yang menunjukkan bahwa kondisi postur tidak terdeteksi. Karena keputusan dilakukan secara independen pada setiap label, satu citra dapat menghasilkan lebih dari satu label positif secara bersamaan.

Metrik evaluasi yang digunakan dalam penelitian ini meliputi:

- **Macro F1-score** digunakan sebagai metrik utama karena menghitung nilai *F1-score* pada setiap label secara independen, kemudian mengambil nilai rata-ratanya. Metrik ini penting untuk mengevaluasi dataset yang tidak seimbang karena memberikan bobot yang sama terhadap setiap label, termasuk label minoritas seperti PK. Nilai *Macro F1-score* yang lebih tinggi menunjukkan bahwa model memiliki performa yang lebih konsisten pada seluruh label.
- **Micro F1-score** digunakan sebagai metrik sekunder untuk mengukur performa model secara keseluruhan. Perhitungannya dilakukan dengan mengakumulasi nilai *true positive*, *false positive*, dan *false negative* dari seluruh label sebelum menghitung nilai *F1-score*. Metrik ini memberikan gambaran mengenai kemampuan prediksi model berdasarkan seluruh keputusan klasifikasi yang dihasilkan.
- **Hamming Loss** digunakan untuk mengukur proporsi kesalahan prediksi terhadap keseluruhan label yang dievaluasi. Kesalahan tersebut mencakup label positif yang tidak berhasil dideteksi maupun label negatif yang salah diprediksi sebagai positif. Semakin kecil nilai *Hamming Loss*, semakin baik kemampuan model dalam memprediksi setiap label secara individual.

Selain ketiga metrik utama tersebut, nilai *precision*, *recall*, dan *F1-score* pada setiap label juga dianalisis untuk mengetahui kemampuan masing-masing model dalam mengenali FHP, PK, RSP, dan NP. Analisis per label digunakan untuk mengidentifikasi kondisi postur yang dapat dikenali dengan baik serta label yang masih sulit diprediksi oleh model.

3. HASIL DAN DISKUSI

Bagian ini menyajikan hasil eksperimen dari sistem klasifikasi *multi-label* permasalahan postur tubuh yang dikembangkan menggunakan fitur landmark MediaPipe Pose. Pembahasan dilakukan secara bertahap, dimulai dari hasil validasi *dataset*, distribusi label setelah pembagian data, hasil seleksi fitur, hingga perbandingan performa model LightGBM, SVM, dan ANN. Evaluasi dilakukan menggunakan metrik *Macro F1-score*, *Micro F1-score*, dan *Hamming Loss* untuk melihat kemampuan model dalam mengenali setiap kategori postur, baik pada kelas mayoritas maupun kelas minoritas. Selain menampilkan hasil kuantitatif, bagian ini juga membahas kecenderungan performa masing-masing model dalam konteks klasifikasi *multi-label* postur tubuh.

Mengacu pada proses pembagian data yang telah dijelaskan pada **Subbab 2.6**, analisis awal pada bagian ini difokuskan pada distribusi label yang terbentuk pada data latih dan data uji. Pemeriksaan distribusi ini penting untuk memastikan bahwa setiap label postur tetap terwakili secara proporsional dalam skenario klasifikasi *multi-label*. Distribusi instans label hasil pembagian data disajikan pada Tabel 6.

Tabel 6. Distribusi Label *Dataset* Setelah *Stratified Split*

Label Kondisi	Data Latih	Data Uji
FHP	521	132
PK	229	57
RSP	299	75
NP	181	45
Total Instans	1230	309

Hasil pembagian data menunjukkan bahwa komposisi data latih dan data uji tetap mewakili kombinasi label pada skenario klasifikasi *multi-label*. Hal ini penting karena distribusi kelas tidak seimbang, terutama pada kelas PK yang jumlah instansnya lebih sedikit. Augmentasi hanya diterapkan pada data latih melalui transformasi koordinat landmark, seperti *Gaussian noise*, *scale up*, *scale down*, dan translasi acak.

Data uji tidak diaugmentasi agar tetap merepresentasikan data baru yang belum pernah dilihat model. Seleksi fitur berhasil mereduksi 186 dimensi awal menjadi 60 fitur final. Fitur yang kurang relevan untuk analisis citra 2D, seperti koordinat *z* dan nilai *visibility* tertentu, disaring terlebih dahulu. Beberapa fitur terkait Kyphosis tetap dipertahankan untuk menjaga sensitivitas model terhadap kelas minoritas.

3.1 Hasil Pengujian dan Komparasi Model

Tiga model *machine learning*, yaitu LightGBM, SVM, dan ANN, dilatih menggunakan 60 fitur spasial terpilih. Pengujian dilakukan secara *multi-label* menggunakan 177 data uji yang tidak pernah dilihat model sebelumnya. Tabel 7 menunjukkan hasil komparasi performa dari ketiga model tersebut.

Tabel 7. Hasil Evaluasi Komparasi Model Multi-Label

Model	F1 Macro	F1 Micro	Hamming Loss
SVM	0,76	0,79	0,19
ANN	0,74	0,78	0,20
LightGBM	0,72	0,77	0,19

Berdasarkan Tabel 7, SVM memperoleh performa keseluruhan terbaik dengan nilai *Macro F1-score* sebesar 0,76 dan *Micro F1-score* sebesar 0,79. Nilai *Hamming Loss* SVM sebesar 0,19, setara dengan LightGBM dan lebih rendah dibandingkan ANN yang memperoleh nilai sebesar 0,20. Karena *Macro F1-score* memberikan bobot yang setara terhadap setiap label, metrik tersebut digunakan sebagai dasar utama dalam menentukan model terbaik pada dataset dengan distribusi label yang tidak seimbang. Hasil ini menunjukkan bahwa SVM memiliki performa yang lebih konsisten dalam mengenali seluruh kondisi postur dibandingkan dua model lainnya.

Keunggulan SVM pada penelitian ini diduga berkaitan dengan karakteristik 60 fitur spasial yang berbentuk numerik kontinu dan memiliki hubungan nonlinier. Kernel Radial Basis Function (RBF) memungkinkan SVM membentuk batas keputusan nonlinier berdasarkan tingkat kemiripan atau jarak antarsampel dalam ruang fitur. Mekanisme maximum margin juga membantu model memperoleh batas pemisah yang lebih optimal, sedangkan penerapan class weighting berkontribusi dalam menjaga keseimbangan performa pada label dengan jumlah data yang berbeda.

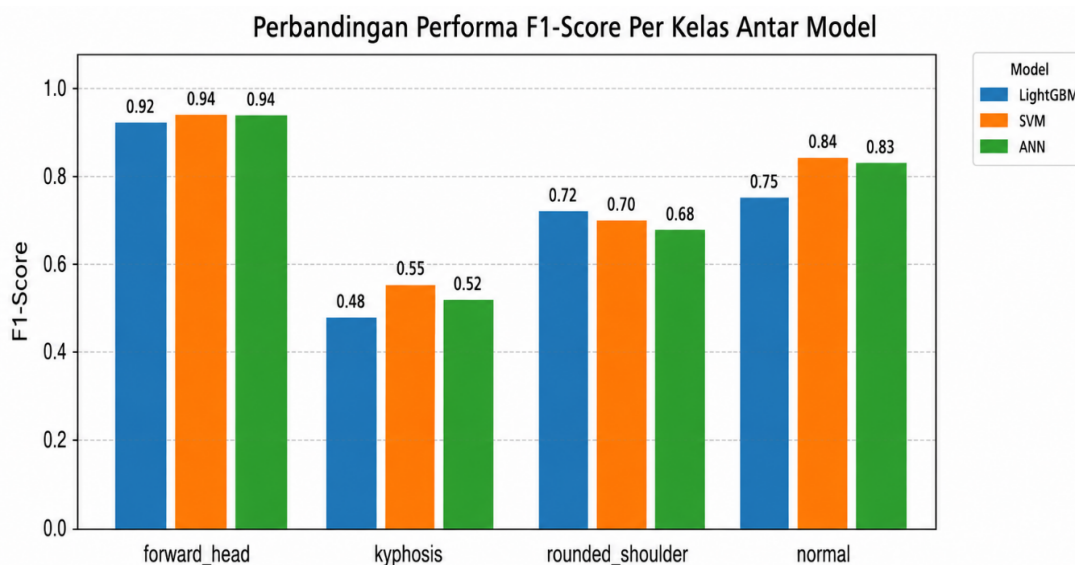
Sebaliknya, LightGBM membentuk keputusan melalui serangkaian pemisahan berbasis pohon pada fitur tertentu. Karakteristik tersebut kemungkinan memberikan keuntungan pada label yang memiliki indikator pembeda relatif kuat, seperti FHP dan RSP. Namun, pendekatan berbasis pohon dapat menjadi lebih terbatas ketika pemisahan antarkondisi postur bergantung pada kombinasi yang lebih halus antara fitur sudut, jarak, rasio, pergeseran posisi, dan keselarasan tubuh.

ANN juga mampu memodelkan hubungan nonlinier, tetapi performanya sangat dipengaruhi oleh jumlah dan keragaman data pelatihan, konfigurasi arsitektur, serta proses optimasi parameter. Pada dataset yang relatif terbatas, SVM kemungkinan lebih efektif dalam membentuk batas keputusan yang stabil. Meskipun demikian, penjelasan tersebut merupakan interpretasi terhadap hasil eksperimen dan belum membuktikan hubungan sebab-akibat secara langsung. Konfirmasi lebih lanjut masih memerlukan analisis tambahan, seperti visualisasi distribusi ruang fitur, pengujian sensitivitas hyperparameter, atau eksperimen ablasi.

Tabel 8. *Classification Report* per Kelas untuk Ketiga Model

Kelas	LGB P	LGB R	LGB F1	SVM P	SVM R	SVM F1	ANN P	ANN R	ANN F1
FHP	0,90	0,95	0,92	0,98	0,89	0,94	0,95	0,93	0,94
PK	0,55	0,42	0,48	0,50	0,61	0,55	0,5	0,54	0,52
RSP	0,71	0,73	0,72	0,69	0,71	0,7	0,65	0,72	0,68
NP	0,82	0,69	0,75	0,75	0,96	0,84	0,81	0,84	0,83

Berdasarkan Tabel 8, ketiga model menunjukkan pola performa yang konsisten. FHP menjadi kelas dengan F1-score tertinggi pada seluruh model, yaitu 0,92–0,94, diikuti oleh NP sebesar 0,75–0,84 dan RSP sebesar 0,68–0,72. Sebaliknya, PK menjadi kelas yang paling sulit diklasifikasikan dengan F1-score terendah, yaitu 0,48–0,55. Pola ini menunjukkan bahwa fitur biomekanik yang digunakan lebih representatif untuk mendeteksi FHP, terutama melalui fitur *ear_hip_x_norm*, sedangkan PK masih memiliki keterbatasan karena tidak adanya titik acuan tubuh pada area vertebra torakal di MediaPipe Pose. Perbandingan *F1-score* per kelas antar model divisualisasikan pada Gambar 5.

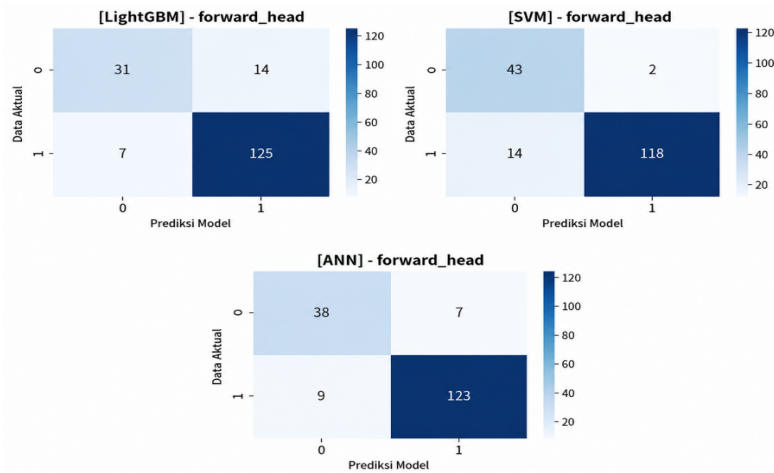


Gambar 5 Perbandingan *F1-score* Per Kelas Antar Model

Berdasarkan visualisasi pada Gambar 5, ketiga model menunjukkan performa yang relatif kompetitif pada setiap kelas postur. Secara umum, SVM menunjukkan kinerja yang paling konsisten, terutama pada kelas PK dan NP. Pada kelas FHP, SVM dan ANN memperoleh *F1-score* tertinggi sebesar 0,94, sedikit lebih tinggi dibandingkan LightGBM sebesar 0,92. Sementara itu, pada kelas PK, SVM juga menghasilkan nilai tertinggi, yaitu 0,55, diikuti oleh ANN sebesar 0,52 dan LightGBM sebesar 0,48. Hal ini menunjukkan bahwa meskipun kelas PK merupakan kelas yang paling sulit diklasifikasikan, SVM masih mampu memberikan performa yang lebih baik dibandingkan dua model lainnya.

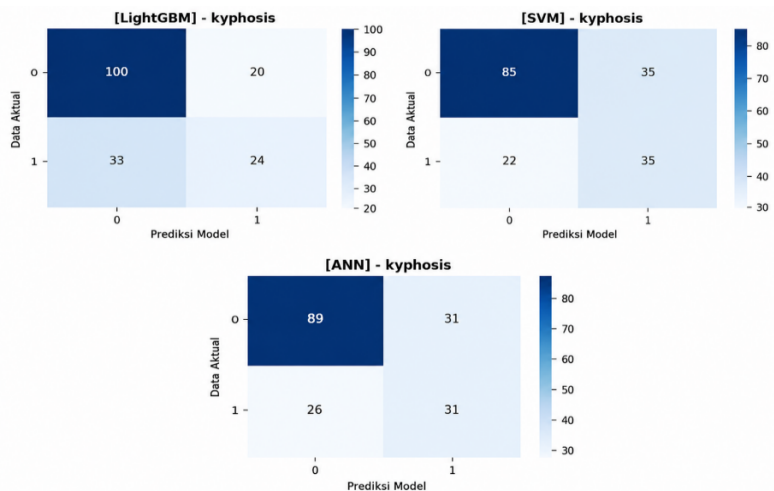
3.2 Analisis Komparatif Confusion Matrix Antar Label

Untuk memperjelas perbandingan performa ketiga model pada tiap kondisi postur, confusion matrix disajikan dalam susunan per-label, di mana setiap label memiliki satu figur berisi tiga confusion matrix LightGBM, SVM, dan ANN yang ditampilkan berdampingan. Susunan ini memungkinkan perbandingan langsung terhadap pola kesalahan klasifikasi (TN, FP, FN, dan TP) ketiga model pada kondisi postur yang sama, tanpa harus membandingkan antar-figur yang terpisah seperti pada susunan per-model sebelumnya.



Gambar 6. FHP Confusion Matrix

Gambar 6 menunjukkan bahwa ketiga model mencapai performa yang sangat tinggi pada kelas FHP, dengan LightGBM mencatatkan TP tertinggi sebesar 125 dan FN terendah sebesar 7 dari total 132 sampel aktual, menghasilkan *Recall* 0,95. SVM justru berada di posisi *Recall* terendah (0,90, FN = 14) namun memiliki FP yang sangat minim yaitu hanya 2, menghasilkan *Precision* tertinggi (0,98) di antara ketiga model. ANN berada di posisi tengah dengan keseimbangan FP = 7 dan FN = 9, menghasilkan *Recall* 0,93. Pola ini mengonfirmasi bahwa fitur proxy CVA (*ear_hip_x_norm*) yang menjadi sinyal utama FHP cukup kuat sehingga ketiga model, meskipun berbeda pendekatan (*ensemble tree*, *margin* maksimum, dan jaringan saraf), sama-sama mampu memisahkan kelas ini secara efektif.



Gambar 7. PK Confusion Matrix

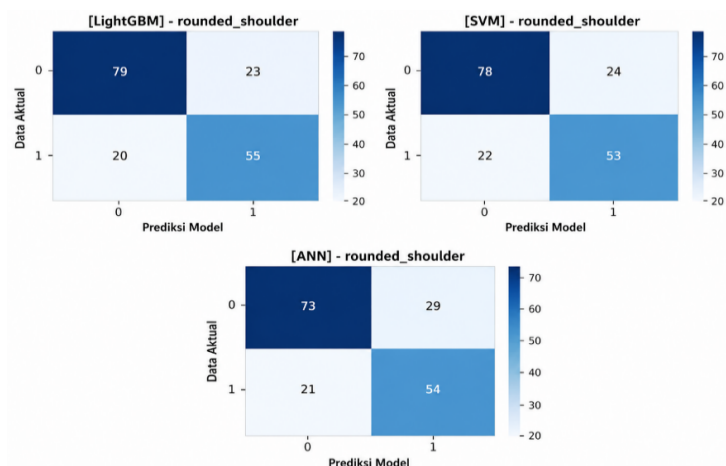
Gambar 7 memperlihatkan bahwa *Postural Kyphosis* (PK) merupakan label dengan tantangan klasifikasi terbesar bagi seluruh model. LightGBM mencatat nilai *recall* terendah, yaitu 0,42 dengan TP = 24 dan FN = 33, meskipun memiliki *precision* sebesar 0,55. Pola tersebut menunjukkan bahwa LightGBM cenderung lebih konservatif dalam memprediksi indikasi PK, yaitu hanya mengenali data dengan pola fitur yang relatif jelas sehingga masih terdapat cukup banyak kasus PK yang tidak terdeteksi.

Sebaliknya, SVM memperoleh *recall* tertinggi sebesar 0,61 dengan TP = 35, tetapi juga menghasilkan jumlah FP tertinggi, yaitu 35. Hasil tersebut menunjukkan bahwa SVM lebih sensitif dalam mendeteksi indikasi PK, tetapi

juga lebih sering mengklasifikasikan data non-PK sebagai PK. Sementara itu, ANN menunjukkan performa yang berada di antara LightGBM dan SVM dengan *recall* sebesar 0,54, FP = 31, dan FN = 26.

Perbedaan pola kesalahan tersebut menunjukkan bahwa tantangan klasifikasi PK tidak hanya berkaitan dengan karakteristik algoritma, tetapi juga dengan keterbatasan representasi fitur. MediaPipe Pose tidak menyediakan *landmark* yang secara langsung merepresentasikan area vertebra torakal, padahal karakteristik utama *Postural Kyphosis* berkaitan dengan peningkatan kelengkungan pada area tersebut. Oleh karena itu, pengenalan indikasi PK masih bergantung pada fitur proksi, seperti *Curvature Index*, kemiringan tubuh, keselarasan bahu, posisi kepala, dan rasio antartitik tubuh.

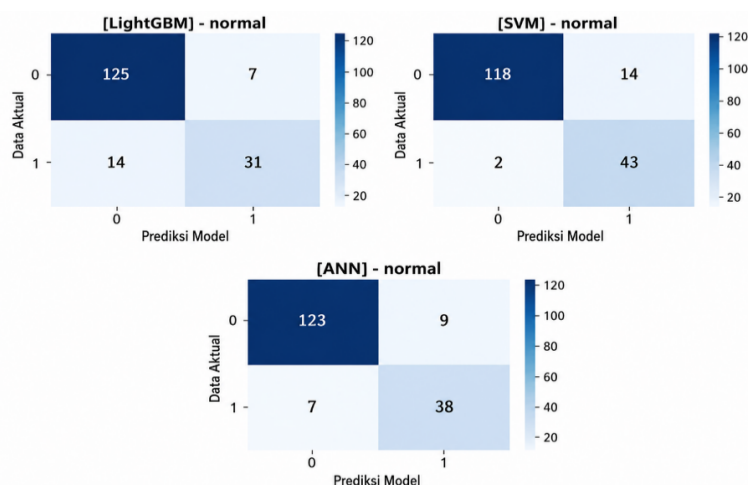
Fitur-fitur tersebut kemungkinan memiliki pola yang tumpang tindih dengan FHP dan RSP sehingga batas antarlabeled menjadi lebih sulit dibedakan. Kondisi ini menyebabkan nilai *F1-score* PK tetap menjadi yang terendah pada seluruh model, meskipun SVM menghasilkan performa yang relatif lebih baik dibandingkan LightGBM dan ANN.



Gambar 8. RSP Confusion Matrix

Gambar 8 mengilustrasikan bahwa performa ketiga model pada kelas RSP relatif berimbang. LightGBM mencatatkan nilai *Recall* tertinggi, yaitu 0,73 dengan TP=55 dan FN=20, diikuti oleh ANN sebesar 0,72 dengan TP=54 dan FN=21, serta SVM sebesar 0,71 dengan TP=53 dan FN=22. Dari sisi FP, LightGBM juga menghasilkan jumlah kesalahan terendah, yaitu 23, dibandingkan SVM sebesar 24 dan ANN sebesar 29. Hasil ini menunjukkan bahwa LightGBM sedikit lebih unggul dalam mendeteksi RSP dibandingkan dua model lainnya.

Meskipun demikian, jumlah FP pada ketiga model masih relatif tinggi, yaitu berada pada rentang 23–29. Pola ini mengindikasikan adanya tumpang tindih fitur (*feature overlap*) antara kelas RSP dan PK. Temuan ini sejalan dengan hasil analisis kesalahan kualitatif yang menunjukkan bahwa kedua kondisi tersebut sering muncul secara bersamaan, sehingga model lebih sulit membedakan batas karakteristik biomekanik antara RSP dan PK.



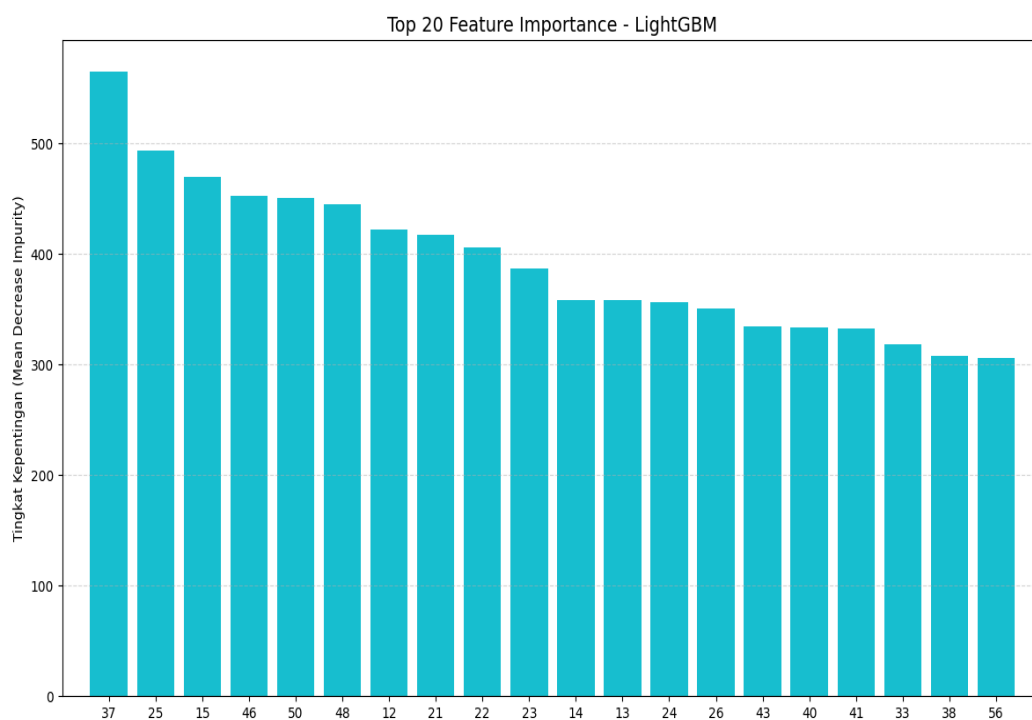
Gambar 9. NP Confusion Matrix

Gambar 9 memperlihatkan bahwa SVM merupakan model yang paling unggul pada kelas NP, dengan nilai *Recall* mencapai 0,96 (TP = 43, FN = 2), lebih tinggi dibandingkan ANN sebesar 0,84 (TP=38, FN=7) dan LightGBM sebesar 0,69 (TP = 31, FN = 14). Hasil ini menunjukkan bahwa SVM paling mampu mengenali subjek yang benar-benar berada pada kondisi normal. Namun, SVM juga mencatat jumlah FP tertinggi, yaitu 14, dibandingkan ANN (9) dan LightGBM (7). Hal ini berarti SVM sedikit lebih sering salah mengklasifikasikan postur bermasalah ringan sebagai normal.

Kondisi tersebut menunjukkan adanya *trade-off* antara kemampuan model dalam mengenali data berlabel NP dan kemampuannya menghindari kesalahan klasifikasi data dengan indikasi masalah postur sebagai kondisi normal. Interpretasi hasil ini ditempatkan dalam konteks sistem skrining awal dan bukan sebagai diagnosis klinis.

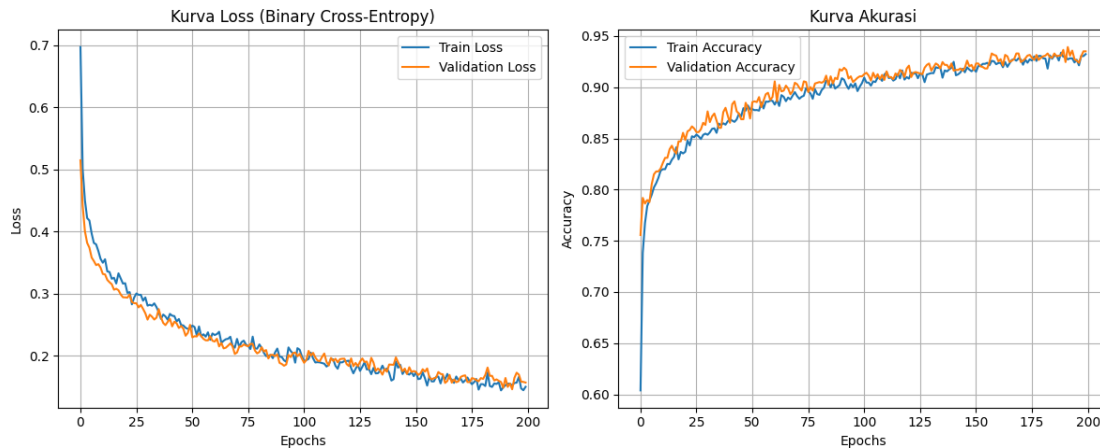
Pada evaluasi label NP, FN menunjukkan data yang sebenarnya berlabel normal, tetapi diprediksi tidak termasuk NP atau diprediksi memiliki indikasi masalah postur. Kesalahan tersebut dapat menyebabkan subjek dengan postur normal memperoleh peringatan yang tidak tepat. Sebaliknya, FP menunjukkan data dengan label kelainan postur yang diprediksi sebagai NP atau kondisi normal. Kesalahan kedua perlu mendapat perhatian karena dapat menyebabkan indikasi postur bermasalah tidak teridentifikasi oleh sistem.

Meskipun SVM memiliki *recall* tertinggi pada label NP, jumlah FP yang lebih besar menunjukkan bahwa peningkatan kemampuan mengenali kondisi normal disertai peningkatan risiko mengklasifikasikan sebagian data dengan indikasi masalah postur sebagai kondisi normal. Oleh karena itu, penerapan model sebagai sistem skrining awal perlu mempertimbangkan keseimbangan antara *recall* NP dan jumlah FP, serta tetap memerlukan verifikasi lebih lanjut oleh tenaga profesional.



Gambar 10. Top 20 Feature Importance LightGBM (Mean Decrease Impurity)

MDI mengukur seberapa besar sebuah fitur berkontribusi dalam mengurangi impuritas pada pembelahan *node* di seluruh estimator LightGBM semakin tinggi nilai MDI, semakin besar kontribusi fitur tersebut. Dari Gambar 10, distribusi *feature importance* relatif merata di antara 20 fitur teratas (MDI berkisar 306–565), tanpa satu fitur yang mendominasi secara ekstrem, meskipun fitur peringkat pertama (indeks fitur 37) sedikit lebih menonjol dibandingkan fitur-fitur lainnya. Hal ini mengindikasikan model memanfaatkan kombinasi beragam fitur secara seimbang, yang merupakan karakteristik yang diinginkan karena menunjukkan robustness model. Fitur-fitur teratas mencakup kombinasi koordinat *raw landmark* dari titik anatomis kritis dan feature engineered dari blok FHP dan Global, memvalidasi bahwa desain *feature engineering* berbasis *proxy* klinis yang diterapkan memberikan kontribusi yang nyata terhadap kemampuan klasifikasi model.



Gambar 11. Learning Curve ANN

Berdasarkan kurva *loss* pada Gambar 11, terlihat bahwa *training loss* dan *validation loss* mengalami penurunan tajam selama 25 *epoch* pertama, yaitu dari sekitar 0,68 menjadi 0,30. Pola tersebut menunjukkan bahwa model mampu mempelajari pola utama pada data secara efektif pada fase awal pelatihan. Setelah itu, nilai *loss* terus menurun secara lebih bertahap hingga *epoch* ke-200, dengan nilai akhir sekitar 0,13 pada *training loss* dan 0,18 pada *validation loss*.

Pola serupa terlihat pada kurva akurasi. *Training accuracy* dan *validation accuracy* meningkat tajam dari kisaran 0,75–0,78 pada *epoch* awal menjadi sekitar 0,90 dalam 75 *epoch* pertama. Selanjutnya, nilai akurasi terus meningkat secara bertahap hingga mencapai kisaran 0,92–0,94 pada *epoch* ke-200. Kurva *validation accuracy* juga berada relatif sejajar dengan kurva *training accuracy* dan pada beberapa bagian menunjukkan nilai yang sedikit lebih tinggi.

Kurva validasi pada Gambar 11 diperoleh dari sebagian data latih yang dipisahkan sebagai data validasi internal, bukan dari 177 citra data uji. Data uji tetap disimpan secara terpisah dan hanya digunakan pada tahap evaluasi akhir model. Kedekatan antara kurva pelatihan dan validasi, baik pada metrik *loss* maupun akurasi, menunjukkan bahwa model ANN memiliki kemampuan generalisasi yang baik dan tidak mengalami *overfitting* yang signifikan selama proses pelatihan. Kondisi tersebut juga menunjukkan bahwa penerapan *Batch Normalization* dan *Dropout* sebesar 0,30 pada setiap *hidden layer* berkontribusi dalam menjaga stabilitas pelatihan dan mengurangi risiko *overfitting*.

4. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan sistem klasifikasi multi-label untuk mengenali indikasi *Forward Head Posture* (FHP), *Postural Kyphosis* (PK), *Rounded Shoulder Posture* (RSP), dan *Normal Posture* (NP) yang dapat berkaitan dengan gaya hidup sedentari dan kebiasaan duduk berkepanjangan melalui analisis citra postur statis dalam posisi berdiri menggunakan MediaPipe Pose dan *machine learning*. Pendekatan *feature engineering* berbasis biomekanik mampu mengubah koordinat *landmark* tubuh menjadi representasi numerik yang lebih informatif, meliputi fitur jarak, sudut, rasio, pergeseran posisi, dan keselarasan tubuh, sehingga beberapa kondisi postur dapat dikenali secara simultan.

Hasil komparasi menunjukkan bahwa SVM memberikan performa keseluruhan terbaik dengan *Macro F1-score* sebesar 0,76 dan *Micro F1-score* sebesar 0,79. Meskipun demikian, LightGBM memperoleh nilai *recall* yang lebih tinggi pada label FHP dan RSP. Sementara itu, PK menjadi label yang paling sulit dikenali karena keterbatasan *landmark* dua dimensi MediaPipe Pose dalam merepresentasikan kelengkungan sagital pada area vertebra torakal.

Secara keseluruhan, kombinasi MediaPipe Pose, *feature engineering*, dan *machine learning* berpotensi digunakan sebagai dasar pengembangan sistem pemantauan indikasi postur secara waktu nyata untuk mendukung proses skrining awal. Sistem yang dikembangkan tidak dimaksudkan sebagai alat diagnosis klinis, sehingga hasil prediksi tetap memerlukan verifikasi lebih lanjut oleh tenaga profesional.

Penelitian selanjutnya dapat meningkatkan representasi PK melalui penggunaan kamera *multiview*, *Inertial Measurement Unit* (IMU), sensor tekanan, atau pendekatan *sensor fusion*. Pengembangan berikutnya juga dapat mengevaluasi metode *ensemble learning* dan *stacking*, melakukan optimasi nilai *threshold* secara khusus untuk setiap label, serta menggunakan anotasi atau validasi dari tenaga profesional untuk meningkatkan reliabilitas hasil klasifikasi.

REFERENSI

- [1] F. Roggio, S. Di Grande, S. Cavalieri, D. Falla, and G. Musumeci, "Biomechanical Posture Analysis in Healthy Adults with Machine Learning: Applicability and Reliability," *Sensors*, vol. 24, no. 9, May 2024, doi: 10.3390/s24092929.
- [2] J. E. Estrada, L. A. Veal, and M. Devaraj, "Modelling Proper and Improper Sitting Posture of Computer Users Using Machine Vision for a Human-Computer Intelligent Interactive System during COVID-19," *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095402.
- [3] G. Çobanoğlu, M. Y. Demirkan, Z. B. Ececi, and N. Atalay Güzel, "Forward Head Posture and Its Effect on Muscle Activation," *Gazi Sağlık Bilimleri Dergisi*, vol. 9, no. 1, pp. 85–93, Mar. 2024, doi: 10.52881/gsbdergi.1376080.
- [4] S. Tavakoli and A. Maleki, "Proposing Myoelectric Biomarkers for Assessing Kyphosis," *InfoScience Trends*, vol. 2, no. 8, pp. 44–56, Aug. 2025, doi: 10.61882/ist.202502.08.04.
- [5] M. Khan *et al.*, "RELATIONSHIP OF ROUNDED SHOULDER POSTURE WITH PEAK EXPIRATORY FLOW RATE IN UNDERGRADUATE MEDICAL STUDENTS HAVING MECHANICAL NECK PAIN," 2024.
- [6] W. Abu-Taleb, A. A. Yamany, Y. M. Aneis, and S. T. Abu El Kasem, "Effect of adding global postural reeducation to kendall exercises for treating asymptomatic forward head posture: A single-blinded randomized controlled trial," *J. Bodyw. Mov. Ther.*, vol. 42, pp. 938–947, Jun. 2025, doi: 10.1016/j.jbmt.2025.01.039.
- [7] F. Roggio, B. Trovato, M. Sortino, and G. Musumeci, "A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: A narrative review," Nov. 15, 2024, *Elsevier Ltd.* doi: 10.1016/j.heliyon.2024.e39977.
- [8] M. Menanno, C. Riccio, V. Benedetto, F. Gissi, M. M. Savino, and L. Troiano, "An Ergonomic Risk Assessment System Based on 3D Human Pose Estimation and Collaborative Robot," *Applied Sciences (Switzerland)*, vol. 14, no. 11, Jun. 2024, doi: 10.3390/app14114823.
- [9] I. Nurdiansyah, R. Utami, and ; Muchamad Sandy, "Machine Learning System untuk Mendeteksi Gerakan Tubuh Menggunakan Library Mediapipe," Jan. 2025. doi: <https://doi.org/10.22441/format.2025.v14.i1.008>.
- [10] P. Kongchansawang, T. Naowavathong, R. Srisuttee, and S. Moodleah, "Automated Classification of User Exercise Poses in Virtual Reality Using Machine Learning-Based Human Pose Estimation," *IEEE Access*, vol. 13, pp. 179789–179802, 2025, doi: 10.1109/ACCESS.2025.3622115.
- [11] J. H. Meghana, H. K. Chethan, K. S. Santhosh Kumar, and S. P. S. Prakash, "Comprehensive Analysis of Pose Estimation and Machine Learning Classifiers for Precise Yoga Pose Detection and Classification," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 3345–3356. doi: 10.1016/j.procs.2025.04.592.
- [12] S. Takigami *et al.*, "Estimation of Shoulder Joint Rotation Angle Using Tablet Device and Pose Estimation Artificial Intelligence Model," *Sensors*, vol. 24, no. 9, May 2024, doi: 10.3390/s24092912.
- [13] J. Shen, L. Chen, X. He, C. Zuo, X. Li, and L. Dong, "An Interactive Human-in-the-Loop Framework for Skeleton-Based Posture Recognition in Model Education," *Biomimetics*, vol. 10, no. 7, Jul. 2025, doi: 10.3390/biomimetics10070431.
- [14] H. Zhang, D. Gračanin, W. Zhou, D. Dudash, and G. Rushton, "Toward Real-Time Posture Classification: Reality Check," *Electronics (Switzerland)*, vol. 14, no. 9, May 2025, doi: 10.3390/electronics14091876.
- [15] A. Tharatipyakul, T. Srikaewsiew, and S. Pongnumkul, "Deep learning-based human body pose estimation in providing feedback for physical movement: A review," Sep. 15, 2024, *Elsevier Ltd.* doi: 10.1016/j.heliyon.2024.e36589.
- [16] A. Peryanto, A. Yudhana, and D. R. Umar, "Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network," 2019. [Online]. Available: <https://www.mathworks.com/discovery/convolutional-neural-network.html>
- [17] Y. Liu *et al.*, "Representation-centric survey of supervised skeletal action recognition and the new benchmark," *Pattern Recognit.*, vol. 180, p. 114140, Dec. 2026, doi: 10.1016/j.patcog.2026.114140.
- [18] M. R. Abrori, F. T. Anggraeny, and E. Y. Puspaningrum, "Analisis Fitur Pada Citra Gestur Tangan Sistem Isyarat Bahasa Indonesia," *FORMAT: Jurnal Ilmiah Teknik Informatika*, vol. 12, no. 2, pp. 165–172, Dec. 2023.
- [19] I. H. Al Ghazali, M. A. Fathin, and A. R. Handoko, "A Comparative Study of Machine Learning with Statistical Feature Selection for Risk Detection of Diabetic," *Jurnal Ilmiah FIFO*, vol. 17, no. 2, p. 102, Nov. 2025, doi: 10.22441/fifo.2025.v17i2.001.
- [20] T. H. Pinem and Z. P. Putra, "Evaluasi Kinerja Algoritma Klasifikasi Deep Learning dalam Prediksi Diabetes," *Jurnal Ilmiah FIFO*, vol. 17, no. 1, p. 17, Jun. 2025, doi: 10.22441/fifo.2025.v17i1.003.